Assessment and comparison of different approaches for mask write time reduction

A. Elayat[1], T. Lin[2], E. Sahouria[2], S. F. Schulze[1]

[1]Mentor Graphics Corp. 8005 SW Boeckman Rd, Wilsonville, OR 97070
[2]Mentor Graphics Corp. 46871 Bayside Parkway, Fremont, CA 94538

ABSTRACT

The extension of 193nm exposure wavelength to smaller nodes continues the trend of increased data complexity and subsequently longer mask writing times. We review the data preparation steps post tapeout, how they influence shot count as the main driver for mask writing time and techniques to reduce that impact. The paper discusses the application of resolution enhancements and layout simplification techniques; the fracture step and optimization methods; mask writing and novel ideas for shot count reduction.

The paper will describe and compare the following techniques: optimized fracture, pre-fracture jog alignment, generalization of shot definition (L-shot), multi-resolution writing, optimized-based fracture, and optimized OPC output. The comparison of shot count reduction techniques will consider the impact of changes to the current state of the art using the following criteria: computational effort, CD control on the mask, mask rule compliance for manufacturing and inspection, and the software and hardware changes required to achieve the mask write time reduction. The paper will introduce the concepts and present some data preparation results based on process correction and fracturing tools.

Keywords: mask write time reduction, shot count, jog alignment, L-shot, multi-resolution writing, optimized-based fracture

1. INTRODUCTION

In traditional fracture, primitive shapes are created to partition the complex input polygons submitted to the fracture algorithm. The post-OPC layout must be represented by combining these trapezoids of various sizes and configurations. Mask write tools then form mask patterns through the sequential exposure of basic trapezoidal shapes (usually rectangular and right-angled triangular shots), where the tool decomposes the post-fracture shapes further into exposure shots. The main drivers for mask cost have been mask lithography tools, which comprise the most expensive mask manufacturing equipment. The time required for a mask writing machine to expose a mask defines the contribution of equipment depreciation and maintenance cost to an individual mask. The number of shots directly correlates to a large majority of the mask write time. To contain mask cost, shot count needs to be minimized. [1]

As designs get smaller and more complex, aggressive OPC treatments that result in highly fragmented layouts are applied, steadily increasing the total shot count for an advanced photomask. The introduction of inverse lithography OPC solutions that output curvilinear or "raw" masks further accelerates this trend. Here the number of trapezoids needed to approximate a curve becomes too large to be practically written using vector-shaped beam writers. Hence simplifications are required to make the mask manufacturable.

This paper introduces and describes shot count reduction strategies and assesses the benefits and challenges associated with these mask write time reduction techniques.

2. SHOT COUNT REDUCTION STRATEGIES

2.1 Optimized fracture

Conventional fracture is a necessary step in any mask preparation flow. It converts polygon data into numerical control data in a format appropriate to the mask writing and inspection tools. Advanced fracturing requires additional techniques to be used in conjunction with fracture. These are the ability to model and simulate mask images. Since the fracture step creates the trapezoidal representation of the data that a machine transforms further into the exposure shots – optimizing this step is the first and logical starting point for shot count reduction. In a recent experiment the Calibre fracture engine was tuned for the 20/22nm technology node. In tests on an M1 22nm design, roughly a 2% shot count reduction was seen when comparing software versions before and after the latest fracture engine enhancement. While the reduction is not large in itself, such improvements have a large cumulative effect over time; and other algorithm improvements, such as small-outside figure reduction, also indirectly improve the shot count. [2] More importantly, the core fracture engine itself is a platform for additional techniques that produce much more significant shot count reduction benefits, as will be shown in the following sections.

2.2 Jog alignment (MASKopt)

MASKopt is a Calibre tool that performs jog alignment as the last data processing step prior to fracture. [3] It aligns jogs within user-specified parameters to eliminate small shots between misaligned jogs. Misaligned jogs can occur in OPC during fragmentation when different data levels are merged prior to fracture, or during biasing. When jogs are misaligned even by a small amount, a small trapezoid is required between them.
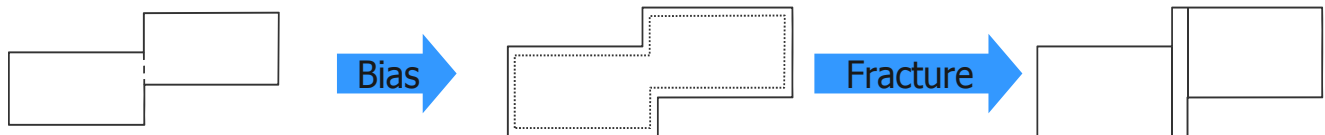

Figure 1. Misalignment of jogs during the application of a process bias.

In the ideal case, the data provided to the mask house for writing would have jogs aligned wherever possible. Even so, the bias required to compensate for the mask process will cause the jogs to become misaligned again. Figure 1 shows how applying a bias causes jogs to become misaligned.

Jog alignment as conducted by MASKopt provides a way to realign the jogs and avoid creating many small trapezoids. The principle is shown in Figure 2. Comparing the fracture without MASKopt in the top row, with the fracture after MASKopt in the bottom row, it can be seen that MASKopt has eliminated the small sliver shot in the center of the figure.
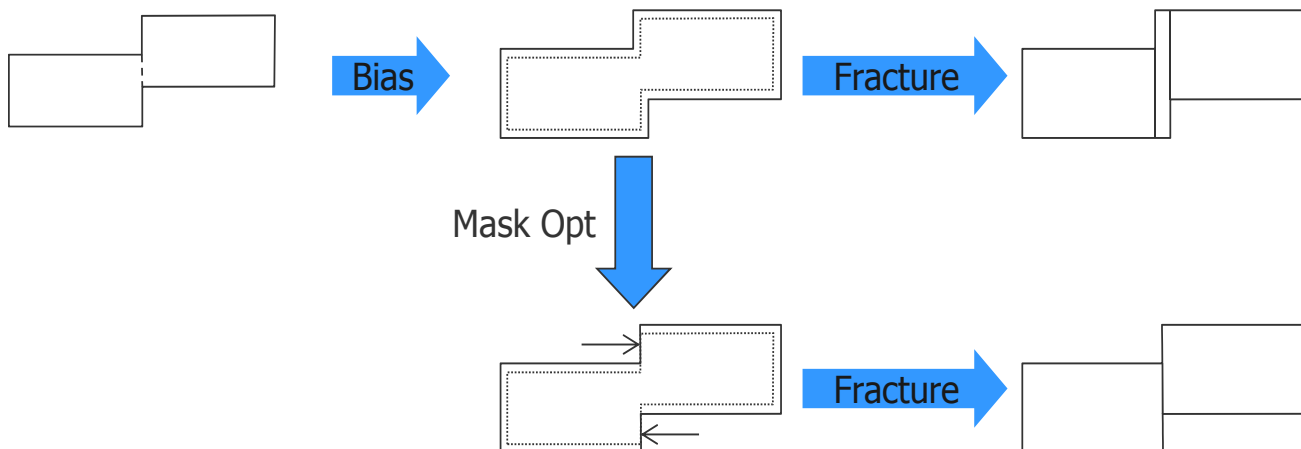

Figure 2. Reduction of shot count by jog alignment using MASKopt

The jog alignment processing step in MASKopt works on OASIS input data and hence, in a manufacturing flow, it can use the same fracture engine as is used today. It is also fully compatible with vector-shaped beam mask writers and their onboard proximity effect correction (PEC) algorithms.

To verify the output jog alignment, MASKopt should be applied only to write data and not to inspection data. This way it can be ensured that the small jog alignments applied only introduce insignificant changes to the mask data and that the results are kept below the detectability limit of the inspection tools.

Jog alignment can yield significant shot count reduction. A 22nm active layer was treated with the Calibre MASKopt tool. The results are shown in in Figure 3. One important metric for any shot count reduction technique is mask error as measured by edge placement error (EPE). A 34% shot count reduction was achieved without any degradation of the mask (based on the EPE range). No attempt in the OPC process was made to align jogs or smooth fragments, which resulted in many small misalignments. Jog alignment is specifically designed to target those situations.
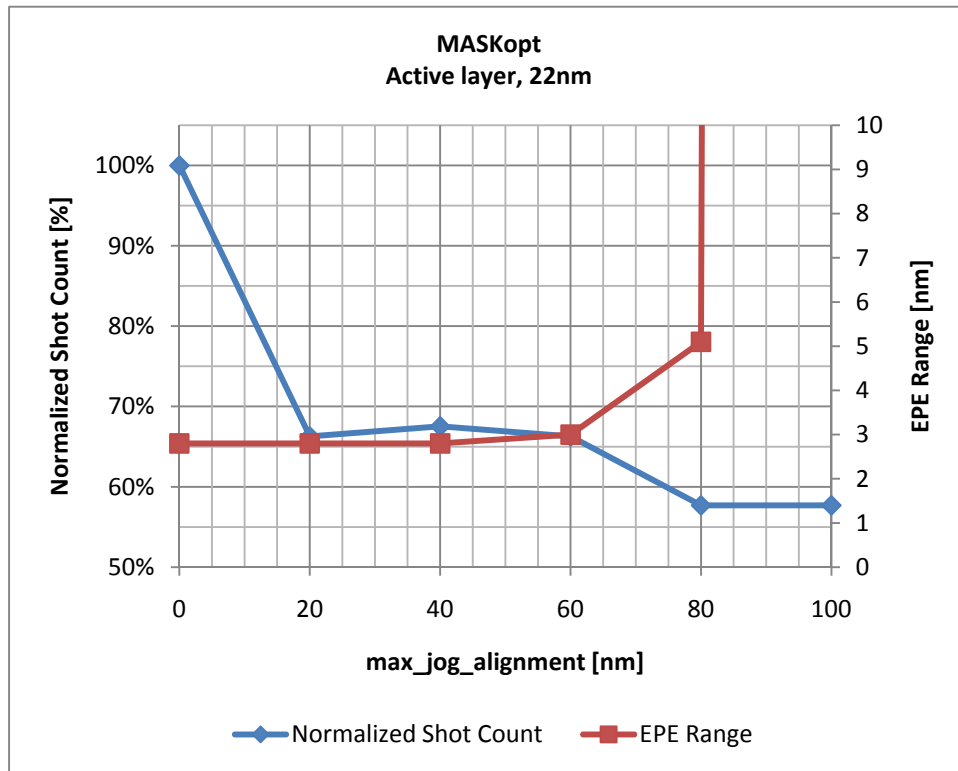

Figure 3. MASKopt results showing shot count and EPE range versus max_jog_alignement at mask scale.

The extent of jog alignment is controlled by three main parameters that the user can specify; the maximum height of a jog that can move, the maximum distance a jog can move, and a spacing parameter to maintain a mask rule constraint for the distance of adjacent jogs on one side of a polygon. In the experiment described in Figure 3, the maximum jog movement distance was varied over a range of 0nm up to 100nm at mask scale. The application of MASKopt should be verified under wafer lithography conditions to ensure that changes to the mask do not adversely impact wafer print quality. Parameters can be fine-tuned with the OPC team to maximize the shot savings while minimizing the contour shift impacting the results of wafer printing. These studies can be conducted with OPC verification tools like Calibre OPCverify.

2.3 L-shot

L-shot fracture reduces shot count by expanding the range of geometries that can be written in a single shot. [4]   Current e-beam mask writing tools allow triangles or rectangles to be written in a single shot. The concept of L-shot fracture is to increase the capability of the write tools so that a single shot in the shape of an "L" may also be written. These L shapes would be obtained by combining two adjacent rectangular shots into a single L thus reducing total shot count – theoretically up to 50%.

L-shaped shots can be employed where only one side of a shape has a jog. This is a very common occurrence. In such situations, traditional fracture would create two rectangles in the shape of an L. With L-shot fracturing, only a single L-shaped shot is created. An example of such a case is shown in Figure 4. Only two shots are required. (The same figure fractured in the traditional manner is shown in Figure 1 and requires 3 shots.)
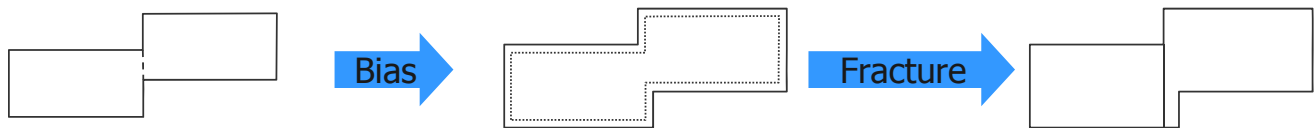

Figure 4. L-shot fracture example

In order to create an L-shaped shot, an additional aperture is required in the write tools. Today two rectangular apertures are used to create rectangular shots of different sizes. In order to create an L-shaped shot, a cross shaped aperture is required. This is shown in Figure 5. Because of this additional aperture, the implementation of L-shot fracture requires additional development by the e-beam write tool manufacturers.
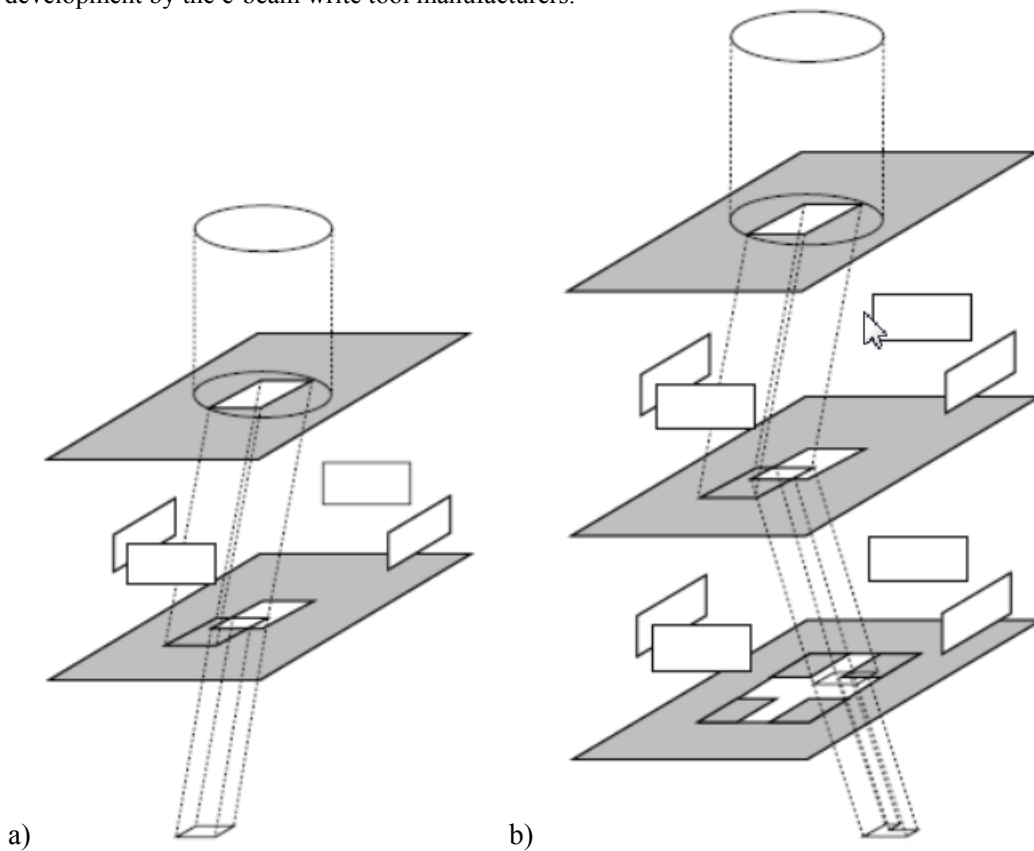

a)   b)

Figure 5. Beam shaping apertures required for a) traditional and b) L-shot strategies [4]

No difficulties with PEC correction are anticipated since the shots do not overlap. The limiting factor in the implementation of this technology is the development of a write tool with the additional aperture required to create L-shaped shots and the respective changes to data format and mask writer software to support it.

L-shot reduces shot count by combining two adjacent rectangles into a single "L"-shaped shot. As in conventional fracture, the L-shot algorithm attempts to avoid the creation of small figures, or slivers. For instance, the dimensions of the "L" are constrained such that no leg is narrow and thin. Furthermore, the entire L-shot is constrained to fit within the specified square shot of the mask writer. [4] In some cases, the exposure avoids a small figure, possibly producing a more accurate mask image. Table 1 shows the shot count reduction for 7 test cases. Overall a shot count reduction of between 20% and 40% is achieved.

SMO clips (22nm)	Shot Count Reduction
TC 1	21%
TC 2	36%
TC 3	29%
TC 4	21%
TC 5	25%
TC 6	29%
TC 7	39%

Table 1. L-shot results for 7 test cases showing the shot count reduction that can be achieved.

L-shaped shots require a significant amount of innovation on the e-beam tool vendors' part. The advantage of L-shaped shots is that the probability of two adjacent rectangles forming an L-shape is about 1 in 3 (based on this data).

2.4 Multi-resolution writing

Today, photomasks are conventionally written using two passes and the shots written in those two passes are identical. The mask write time proportionally increases with each identical pass. However, using identical passes is driven more by practical than fundamental concerns. No special computation is required to customize the patterns for the different passes, and required processing such as on-writer PEC can be reused between the passes. [5]

The objective of multi-resolution writing (MRW) is to jointly customize the shot patterns in both passes. This occurs by exposing one "detail" pass with about as many shots as the conventional pass and one "coarse" pass with much fewer shots. The coarse pass is much simpler than the original pass and the detail pass refines the coarse image such that the final desired image is obtained. The two patterns must be defined jointly with the use of an MPC tool and subject to certain constraints relating to the manufacturing process window.
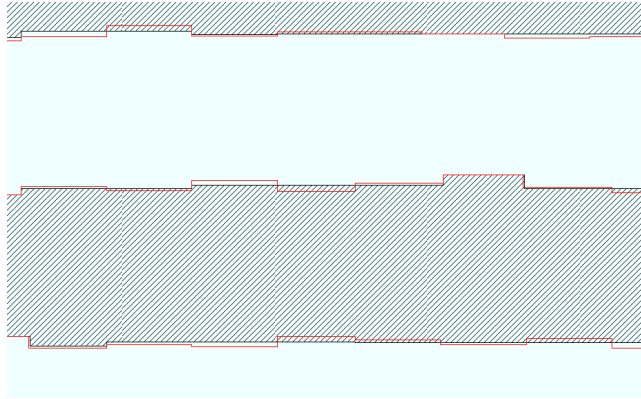
Figure 6. Example of multi-resolution writing. The coarse and detail layers are both shown. [5]

Pattern data has to be provided to the tool in two layers. In OASIS.MASK, each pass is assigned a datatype in the OASIS.MASK format. Another way is to have the two layers overlaid inside of a job deck. Recipe settings are then applied to each layer to set the correct dose and pattern shift. A deeper integration in today's mask writing tools requires machine updates.

Under this framework, PEC is correctly handled by the mask write tools. The tool recognizes the passes in their respective groupings and calculates the pattern density and nominal dose for that group only. In addition to PEC, state of the art mask writing machines correct for fogging effects (exposure related effect) and loading effects (develop and etch process related effects). The concepts of capturing the behavior of fogging and describing fogging effect correction (FEC) are similar to PEC. However, practical considerations result in the following differences:

1) The very large interaction scales make the calculation of FEC by layer likely unnecessary. For the magnitude of the separation used in MRW, the layers are practically identical for the purposes of the FEC kernel.

2) FEC is calculated and stored before commencing the write step. The software will need to correctly discern that the two passes represented by overlaid layers do not constitute a different chip such that it does not "double-count" the contribution of the pass layers.

Loading effect correction (LEC) will behave in the same way as FEC.

A 22nm active layer was treated with the Calibre MRW prototype. The results are shown in in Figure 7. The maxDist parameter controls the aggressiveness of MRW. In this experiment, a sweep of the maxDist parameter was conducted from 0nm to 25nm at mask scale. Shot count reduction of 33% has been achieved by model-based MRW. A noticeable reduction of the EPE range is seen in Figure 7 for the cases where MRW was applied. Compared to the maxDist = 0 nm reference case where MRW was not applied, the EPE range drops when model-based MRW is applied. This is a testament to the effectiveness of model-based mask process correction. Model-based MRW attempts to match the input polygons exactly.
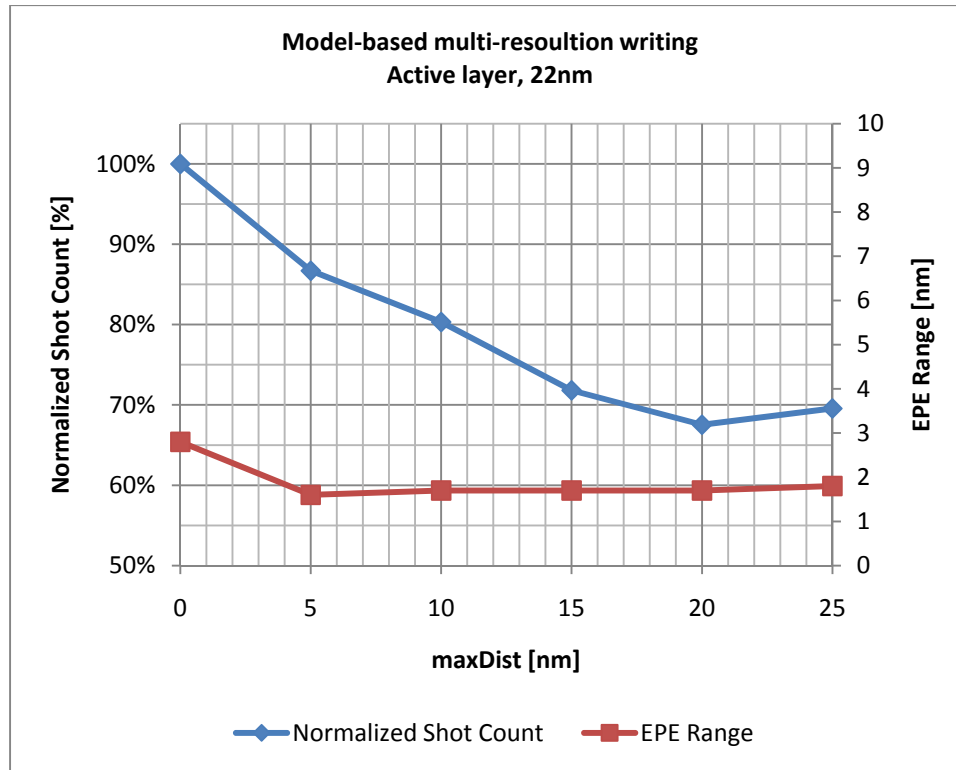
Figure 7. MRW results showing shot count and EPE range versus maxDist at mask scale.

2.5 Optimized-based fracture

Optimized-based fracture (OBF) is significantly different from traditional fracture. In traditional fracture, trapezoids are created to exactly cover the input polygons submitted to the fracture algorithm; shots are abutting and non-overlapping. OBF relaxes those constraints. Shots can be placed such that they overlap or be non-abutting so that sub-resolution gaps exist. The optimization problem is formulated to minimize the number of shots while maintaining the intended post-OPC pattern on the mask. The solution incorporates an e-beam blur (forward scattering + resist blur) model to properly simulate the overlapping and non-abutting shapes. Allowing for overlapping shots and non-abutting shots expands the solution space and provides the optimization engine more opportunity to reduce the shot count. [6]
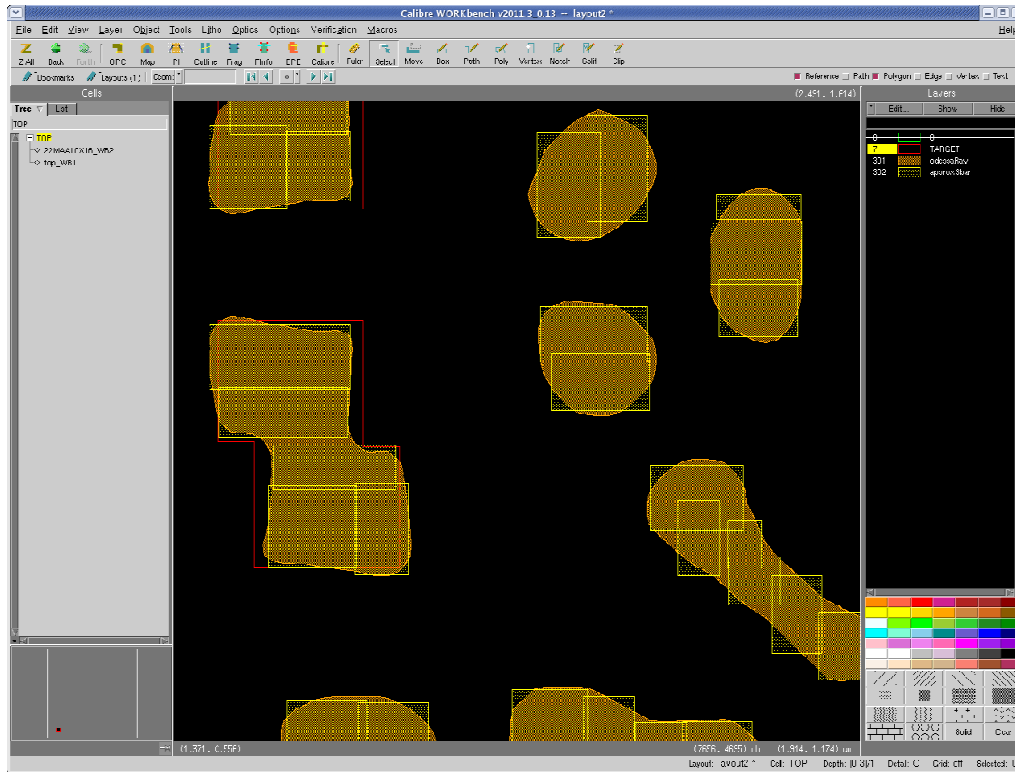
Figure 8. Example of optimization-based fracture on curvilinear masks

OBF has clear advantages when representing curvilinear masks as generated by inverse lithography techniques. It takes advantage of the fact that mask shapes are already rounded and can match this to a simulated e-beam contour. For Manhattan masks, the post-OPC input must be modified slightly to account for beam blur. A subtle point to be made here is that the less blur that is present in the e-beam process, the less opportunity there is to reduce shot count near and around corners, assuming that the desired mask target has to be met to satisfy lithography entitlement.

Special attention must be paid to the PEC, FEC and LEC algorithms when overlapping shots are used. For traditional fracture, the total area of the shots is equal to the pattern area. In the case of overlapping shots, the total shot area is greater than the total area of the pattern. PEC and FEC can correctly treat overlapping shots as the contribution of each shot to backscattered dose is considered regardless of overlap. [7] LEC requires the pattern area of the printed mask shapes to properly calculate the density necessary to correct for downstream processes. This will require an update to the workflow on the current mask writer equipment.

On Manhattan masks, OBF has yielded shot count reduction in the range of 0-20%. On curvilinear or "raw" masks, up to 30% shot count reduction is produced. [8] Another experiment that introduced dose modulation as an additional optimization parameter was done. A very slight additional shot count reduction (0.2%) was observed. Optimization of dose, overlap, shape and position of the trapezoids is a complex problem. [9] The extent of the shot count reduction is a function of the degree of smoothing that can be allowed. This limit necessitates close collaboration between OPC and the mask to determine actual tolerances and identify simplification opportunities.

2.6 Optimized OPC output

The complexity of the OASIS layout presented to the mask manufacturing process is largely driven by the RET and OPC processes. The insertion of assist features, the decoration of layout shapes and the simplification of smooth target mask contours as obtained by inverse lithography methods with tight tolerances increase the vertex count of the output. A number of techniques to reduce the complexity and hence reduce the mask writing time can be applied during the application of OPC. In this case any changes to the output layout are intrinsically verified against the tolerances required

by the litho process. The inverse lithography function of Calibre explicitly models and reduces shot count during the "Manhattanizing step." Calibre OPCpro and Calibre nmOPC tools offer two main user-controlled options to reduce shot count [3]:

1. Jog-smoothing – the alignment of adjacent fragments to eliminate vertices prior to the final iterations
2. Jog-alignment – vertex alignment across the shapes during the fragmentation step

3.   ASSESSMENT OF MASK WRITE TIME SOLUTIONS

Mask write time reduction techniques that are under consideration for the deployment in a running mask manufacturing line require changes that will impact the current technology, workflows and equipment to varying degrees. We will only review approaches that require accommodations in the EDA environment serving the mask writer. For example, new resist materials and increased beam current would be facilitated by the process development teams and their improvements would be orthogonal to improvements made in EDA tools. The generally desirable direction is to obtain maximum write time reduction at the lowest cost and with the smallest impact to the running operation. Changes to the mask write equipment have the longest lead time since they require hardware updates.

A list of potential changes in the mask manufacturing and the wafer manufacturing on the mask customer side includes the following elements:
-   Impact on wafer printing
-   Mask CD control
-   Data preparation time
-   Mask writer changes – format, scheduling, PEC and hardware updates
-   Added work flow complexity

Table 2 establishes the set of rating criteria for the different cost categories and associates a cost indicator that provides a relative rating of the effort for the implementation and execution of the respective technique. Benefit indicators are associated with the potential for shot count reduction. Table 3 details the ratings against these criteria for the various mask write time reduction approaches. The impact on wafer printing due to mask shape alterations is rated into three categories – none, simulation-based lithography verification required, or full yield verification on the wafer is required. The last one applies to all techniques that alter the mask shapes in a significant form. The severity is judged by the degree of deviation from the current methodology of shot formulation and overlap and the expected degradation of the image slope and hence the response to dose variation during mask writes.

Cost Category	Criteria	Cost Indicator
Data preparation effort	Fracture	1
	Rules based optimization	2
	Model based optimization (1 parm)	4
	Model based optmization (2 param)	6
Wafer printing	No impact - Transparent	1
	Lithography verification is required	2
	Yield verification is required	4
Mask CD impact	None	1
	Limited	2
	Moderate (uncertain)	3
Mask writer - format	No change	1
	New primitives/Dose (one change)	2
	New primitives + dose (two changes)	3
Mask writer - scheduling	None	1
	Impact	2
Mask writer - PEC	None	1
	Limited	2
	Moderate	3
Mask writer - Aperture	None	1
	Change	3
Other factors	None	0
	Debug or inspection	1
	Debug and inspection	2

Shot count reduction potential	Benefit Indicator
5%	1
10%	2
15%	3
20%	4
25%	5
30%	6
35%	7
40%	8
45%	9
50%	10
55%	11
60%	12
65%	13
70%	14
75%	15
80%	16
85%	17
90%	18

Table 2. Explanation of ratings used for cost and benefit indicators.

Method: / Impact on:	Optimized OPC output	Optimized fracture	Maskopt	L-Shot	MRW	Opt. based fracture (Geometric - partial overlap)	Opt. based fracture (Geometric+ Dose - partial overlap)
Data processing time	None	None	Rules based approach	None	Model based - one parameter	Model based - one parameter	Model based - two parameters
	1	1	2	1	4	4	6
Wafer printing	None	None	Limited	None	Limited - impact CD uniformity	Limited - controled by parameter	Limited - controled by parameter
	1	1	2	1	4	4	4
Mask CD	None	None	Limited - 2D pattern fidelity	None	Limited - parameter control (separation)	Limited - parameter control (overlap)	Limited - parameter control (overlap, dose)
	1	1	2	1	2	2	2
Mask writer – format	None	None	None	New format	None	None	New format
	1	1	1	2	1	1	2
Mask writer - scheduler	None	None	None	None	Multi-data set handling	None	None
	1	1	1	1	2	1	1
Mask writer – PEC	None	None	None	None	Limited - density averaging	Limited - LEC with merged density	Moderate - PEC/FEC with dose, LEC with merged density
	1	1	1	1	2	2	3
Mask writer - Aperture	None	None	None	New apertures	None	None	None
	1	1	1	3	1	1	1
Other factors	None	None	New Inspection data flow	New debugging infra-structure	New debugging infra-structure	New debugging infra-structure + inspection data flow	New debugging infra-structure + inspection data flow
	0	0	1	1	1	2	2
Benefit	10-30% Standard OPC; 50-90% Inverse Litho	5-15%	5-20%	40%	5-25%	10-30% over litho simplification	15-35% over litho simplifi-cation
	18	3	4	8	5	6	7

Table 3. Cost and benefit ratings for various mask write time reduction solutions.

The data preparation time will increase with the introduction of additional data processing steps. The impact is rated in the following categories – none, rule based, model-based with one optimization parameter, and model-based with two optimization parameters.

The mask writer changes are divided into a variety of categories. Format changes are rated into three categories – none, one change (either primitives or dose modulation), and two changes (introduction of both new primitives and dose modulation). Some techniques require changing the machine internal work flow scheduler – e.g. the separation of data sets into two individual passes. If so, an extra cost factor is added into the rating.

One additional factor to be considered is the e-beam proximity and process correction software. In some cases it is not affected, in other cases only the e-beam proximity correction needs updating. In the worst case, fogging and loading correction methodology also needs updating. On the hardware side, the introduction of new apertures/deflection systems is considered as a cost increase.

A final category emphasizes additional cost factors stemming from increased complexity in the overall mask flow, including extra effort for mask inspection – e.g. an additional data preparation step or new needs for data verification that has to be executed on the altered mask patterning data.

Table 3 also conducts an assessment of the benefits obtainable by the various known methods. The benefit numbers are reported based on data experiments with in-house test cases as conducted at Mentor Graphics. Rating the various influence factors was done in order to guide the focus for development and experimentation. It needs to be noted that the assessment was done through internal experimental results and using general consideration based on known practices. Detailed assessment in the field is in progress to confirm or possibly correct the assumption made during the assessment.

For the comparison, it was assumed that all suggested methods can succeed in producing manufacturing-grade masks (meeting all requirements in CD control, registration control and verified defectivity) and can on average achieve the benefits reported for mask write time reduction across a variety of designs. The results of the assessment are displayed in Figure 9.
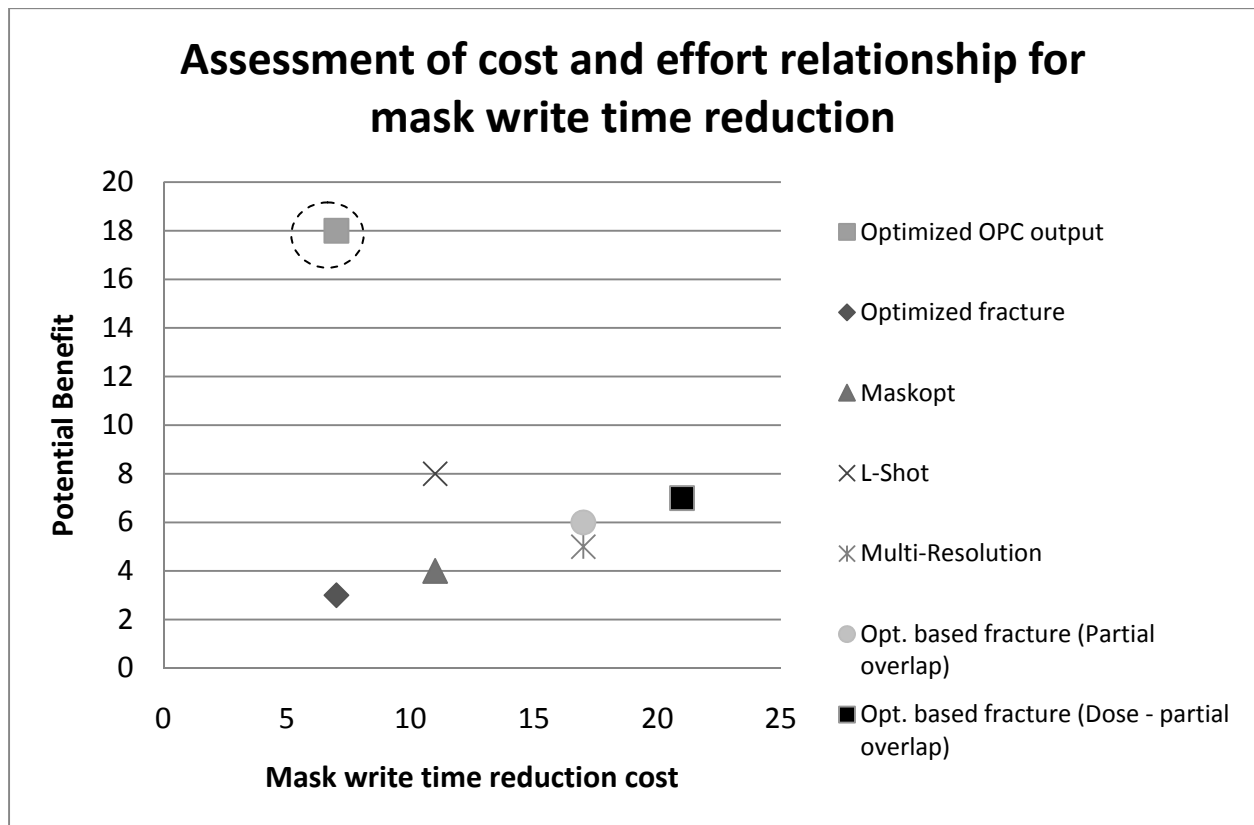

Figure 9. Benefit and effort assessment for various mask write time reduction techniques.

A few observations are noteworthy. The biggest benefit can be obtained by reducing the complexity of the layout through optimization of the OPC (circled above). Integrated verification against target values and tolerances built into advanced OPC prevent the lithographic entitlement from being compromised. Since this is done during OPC, it comes with almost no additional effort during fracture. Optimizations in the fracture tool and simple-rules based improvements like jog alignment (Calibre MASKopt) come next on the effort scale. All methods modifying the mask shapes impose increasing effort depending on the complexity of the changes.

4.   CONCLUSION

In this paper we have reviewed several mask write time reduction techniques with designed to contain the increase in mask shot count while preserving the results quality. It was shown that post-OPC processing techniques can deliver shot count reduction in the range between 20% and 40%. Multiple factors impact the cost associated with shot count reduction – CD control on mask and wafer, hardware and software changes, and data preparation effort. The desirable direction is to obtain maximum write time reduction at the lowest cost and with smallest impact to the running operation. A cost/benefit analysis of the various write time reduction strategies was conducted, which compared the different techniques and illustrated how they rate with one another. While rules-based techniques deliver significant benefit at moderate cost, model-based techniques induce significant cost based on hardware and process changes but provide potentially more accurate mask owing to embedded MPC. The most cost-effective approach is simplification in the OPC stage, where changes to the output layout are intrinsically verified against the tolerances required by the litho-process. As highlighted above, OPC optimization can potentially realize the biggest shot count reduction benefit.

5.   REFERENCES

1.   Chris Spence, Scott Goad, Peter Buck, Richard Gladhill, Russell Cinque, Jürgen Preuninger, Uwe Griesinger, Martin Bloecker, "Mask data volume: historical perspective and future requirements", Proc. SPIE Int. Soc. Opt. Eng. 6281, 62810H (2006)
2.   Steffen Schulze, Emile Sahouria, Eugene Miloslavsky, "High-performance fracturing for variable shaped beam mask writing machines", Proc. SPIE 5130, 648 (2003)
3.   James Word, Keisuke Mizuuchi, Sai Fu, William Brown, Emile Sahouria, "Mask shot count reduction strategies in the OPC flow", Proc. SPIE 7028, 70283F (2008)
4.   Emile Sahouria and Amanda Bowhill, "Generalization of shot definition for variable shaped e-beam machines for write time reduction", Proc. SPIE 7823, 78230T (2010)
5.   Emile Sahouria, "Multiresolution mask writing", Proc. SPIE 7985, 798503 (2011)
6.   Aki Fujimura, Ingo Bork, Taiichi Kiuchi, Tadashi Komagata, Yasutoshi Nakagawa, Kazuyuki Hagiwara, Daisuke Hara, "Improvement of mask write time for curvilinear assist features at 22nm", Proc. SPIE 7823, 782307 (2010)
7.   Christophe Pierrat, Larry Chau, Ingo Bork, "Mask data correction methodology in the context of model-based fracturing and advanced mask models", Proc. SPIE 7973, 79732J (2011)
8.   Timothy Lin, Emile Sahouria, Nataraj Akkiraju, Steffen Schulze, "Reducing shot count through optimization-based fracture", Proc. SPIE 8166 (2011)
9.   Harold R. Zable, Aki Fujimura, Tadashi Komagata, Yasutoshi Nakagawa, John S. Petersen, "Writing wavy metal 1 shapes on 22-nm logic wafers with less shot count", Proc. SPIE 7748, 77480X (2010)
10. Sean Hannon, Travis Lewis, Scott Goad, Kenneth Jantzen, Jianlin Wang, Hien T. Vu, Emile Sahouria, Steffen Schulze, "Reduction of layout complexity for shorter mask write-time", Proc. SPIE 6730, 67303K (2007)
11. Ayman Yehia, "Mask-friendly OPC for a reduced mask cost and writing time", Proc. SPIE 6520, 65203Y (2007)